

# Perbandingan Algoritma Random Forest, Decision Stump, Naïve Bayes, Bayesian Network dan Algoritma C4.5 Untuk Prediksi Pola Kartu Poker

M Fariz Januarsyah  
Magister Teknik Informatika  
Fasilkom, Universitas Sriwijaya  
Palembang, Indonesia  
[m.fariz.januarsyah@gmail.com](mailto:m.fariz.januarsyah@gmail.com)

Ermatita Zuhairi  
Magister Teknik Informatika  
Fasilkom, Universitas Sriwijaya  
Palembang, Indonesia  
[ermatitaz@yahoo.com](mailto:ermatitaz@yahoo.com)

Reza Firsandaya Malik  
Magister Teknik Informatika  
Fasilkom, Universitas Sriwijaya  
Palembang, Indonesia  
[rezaafm@unsri.ac.id](mailto:rezaafm@unsri.ac.id)

**Abstract** - *Poker is one of very popular game in the world with more than hundred million of combinations hand pattern. Due to huge number of combination, poker player is hard to make an accurate decision. The purpose of this research is to give a poker player a suggestion from comparison of different data mining algorithm. This experiment is conducted on C4.5 Algorithm, Decision Stump Algorithm, Naive Bayes Algorithm, Bayesian Network Algorithm, and Random Forest Algorithm on 25.010 data with 11 attribute through ten times of cross-validation.*

**Abstrak**—Poker merupakan salah satu permainan terpopuler didunia dengan kombinasi kartu tangan lebih dari ratusan juta. Karena jumlah yang banyak ini, pemain poker sulit untuk mengambil keputusan yang akurat. Tujuan dari penelitian ini adalah memberikan saran dari perbandingan algoritma data mining yang berbeda. Pengujian algoritma dilakukan pada algoritma C4.5, algoritma Decision Stump, algoritma Naive Bayes, algoritma Bayesian Network, serta algoritma Random Forest dengan menggunakan 25.010 data dengan 11 atribut dan melalui tahap *cross-validation* sebanyak sepuluh (10) kali.

**Kata Kunci**—*Bayesian Network, Algoritma C4.5, Algoritma Naïve Bayes, Random Forest, Decision Stump, Data Mining, Perbandingan Algoritma, Pola Kartu Poker.*

## I. PENDAHULUAN

Perkembangan teknologi informasi pada era globalisasi sekarang berlangsung dengan sangat pesat, hal ini menuntut perkembangan teknologi tersebut untuk dapat saling bersaing. Baik itu dalam kehidupan sehari-hari maupun didalam dunia bisnis dan hiburan yang dinamis dan penuh persaingan. Seiring dengan pertumbuhan bisnis, teknologi tidak lagi bisa unggul secara kompetitif hanya dengan memanfaatkan teknologi yang ada. Keberadaan akan teknologi-teknologi operasional tersebut memang penting, namun kebanyakan dari teknologi tersebut tidak dapat menghasilkan informasi strategis seperti yang diharapkan.

Poker adalah salah satu permainan kartu yang paling populer didunia saat ini [1]. Persaingan yang terjadi dalam dunia permainan poker memaksa para pelakunya untuk selalu memikirkan strategi-strategi dan terobosan yang dapat menjamin kelangsungan prediksi pemainnya [2]. Meskipun peraturan didalam permainan poker itu simpel tetapi permainan poker ini tidak mudah untuk dikuasai [3]. Database dalam jumlah yang besar merupakan salah satu aset berharga yang dimiliki dalam permainan poker. Sebagai salah satu permainan yang sering dimainkan oleh masyarakat luas. Pola Poker yang ada haruslah dapat

digunakan dalam memprediksi strategi untuk mampu menghasilkan suatu informasi yang siap digunakan untuk membantu pengguna dalam mengambil keputusan strategis. Dari data pola kombinasi kartu poker tersebut, pemain poker dapat mengetahui prediksi apa yang akan muncul, seberapa jauh tinggi akurasi prediksi tersebut, dan resiko apa yang harus diambil . Untuk memenuhi kebutuhan-kebutuhan di atas, banyak cara yang dapat ditempuh. Salah satunya adalah dengan melakukan pemanfaatan data base menggunakan teknik *Data Mining* [4].

Bayesian network dan Decision tree adalah salah satu metode yang ada didalam data mining. Penggunaan metode ini telah banyak digunakan didalam penelitian-penelitian yang ada didunia, dari segi bioteknologi penelitian menggunakan metode ini memiliki tingkat akurasi yang tinggi, hingga mencapai 98,51% [5], dari segi medis pun penggunaan metode ini juga mencapai akurasi yang tinggi [6] [7].

Dari penggunaan kedua metode tersebut telah banyak terlahir penelitian-penelitian yang meng-*hybrid* atau menggabungkan kelebihan dari kedua metode itu. Seperti [8] penelitian mereka menggabungkan bayesian clustering dengan decision tree sebagai penggolong untuk mendeteksi gangguan pada sistem. Adapun penelitian lain yang dilakukan oleh [5] penelitian mereka menggabungkan metode bayesian network, principal component analysis, back propagation neural network dan decision tree dalam manajemen pendapatan yang ada pada industri bioteknologi. Sedangkan penelitian dari [9] tentang masalah ketidakseimbangan yang ada pada saat klasifikasi.

Pada pola poker, terdapat banyak klasifikasi yang ada [10]. Dengan jumlah pola yang banyak tersebut maka kemungkinan untuk memprediksi pola yang ada semakin sulit. Adapun penelitian yang dilakukan oleh [4] penelitian tersebut hanya menggunakan satu metode saja yaitu algoritma Bayesian.

Berdasarkan uraian diatas, maka penelitian ini mencoba membandingkan tingkat akurasi dan kecepatan dari algoritma yang berbeda dengan menggunakan data sampel yang sama untuk melihat keunggulan dan kekurangan dari masing-masing algoritma yang berbeda tersebut.

## II. TINJAUAN PUSTAKA

### A. Poker

Poker adalah permainan kartu keluarga yang berbagi taruhan aturan dan biasanya (tapi tidak selalu) dalam peringkat tangan. Permainan Poker berbeda dalam hal bagaimana kartu dibagikan, bagaimana tangan dapat terbentuk, apakah tangan tinggi atau rendah memenangkan taruhan di pertarungan (dalam beberapa Permainan, taruhan dibagi antara tangan tinggi dan rendah), batas taruhan dan bagaimana banyak putaran pertarungan diperbolehkan. Dalam Permainan poker yang paling modern, putaran pertama

pertaruhan dimulai dengan beberapa bentuk kontrak paksa. Aksi kemudian hasil ke kiri. Setiap pemain pada gilirannya baik harus sesuai dengan taruhan sebelumnya maksimum atau lipat, kehilangan jumlah taruhan sejauh ini dan semua bunga lebih lanjut di tangan. Seorang pemain yang cocok dengan taruhan juga dapat meningkatkan, meningkatkan taruhan. Babak taruhan berakhir ketika semua pemain telah baik cocok dengan taruhan terakhir atau dilipat. Jika semua, kecuali satu pemain kali lipat pada setiap putaran, pemain yang tersisa mengumpulkan taruhannya dan dapat memilih untuk menampilkan atau menyembunyikan tangan mereka. Jika lebih dari satu pemain tetap di pertarungan setelah ronde pertarungan final, tangan akan ditampilkan dan tangan memenangkan mengambil taruhannya.

Ada 311.875.200 cara (5-permutasi) untuk lima kartu yang dibagikan dari tumpukan 52 kartu, tapi karena urutan kartu tidak di permasalahan, ada  $5! = 120$  yang diberikan dalam satu sisi tangan, sehingga hanya ada 2.598.960 ( $\frac{52}{5}$ ) kemungkinan kombinasi dalam poker lima kartu. Demikian pula, dalam poker tujuh kartu terdapat 133.784.560 ( $\frac{52}{7}$ ) kombinasi.[4]

### B. Algoritma C4.5

Penggunaan algoritma C4.5 dalam pembuatan aplikasi data mining prediksi kartu poker karena algoritma C4.5 merupakan algoritma yang digunakan untuk melakukan pembentukan pohon keputusan. Pohon keputusan tersebut mampu menghasilkan keputusan kompleks menjadi lebih sederhana, sehingga pengambil keputusan akan lebih menginterpretasikan solusi dari permasalahan.

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut.

- 1) Pilih atribut sebagai akar.
- 2) Buat cabang untuk tiap-tiap nilai.
- 3) Bagi kasus dalam cabang.
- 4) Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain digunakan rumus seperti tertera sebagai berikut :

$$\text{Gain (S, A)} = \text{Entropy(S)} + \sum_{i=1}^n \frac{S_i}{S} + \text{Entropy}(S_i)$$

(Persamaan 1)

Keterangan :

S : Himpunan kasus

A : Atribut

n : Jumlah partisi atribut A

|S<sub>i</sub>| : Jumlah kasus pada partisi ke -i

|S| : Jumlah kasus dalam S

Setelah mendapatkan nilai gain, ada satu hal lagi yang perlu kita lakukan selanjutnya adalah mencari nilai Entropy. Entropy digunakan untuk menentukan seberapa informative sebuah input atribut untuk menghasilkan output atribut.

Rumus dasar dari entropy tersebut adalah sebagai berikut:

$$\text{Entropy (S)} = \sum_{i=1}^n -p_i * \text{Log } p_i$$

(Persamaan 2)

Keterangan :

S : Himpunan kasus

A : Fitur

N : Jumlah partisi S

p<sub>i</sub> : Proporsi dari S<sub>i</sub> terhadap S

### C. Bayesian Network

Naive Bayes merupakan metode yang tidak memiliki aturan, Naive Bayes menggunakan cabang matematika yang dikenal dengan teori probabilitas untuk mencari peluang terbesar dari kemungkinan klasifikasi, dengan cara melihat frekuensi tiap klasifikasi pada data training. Naive Bayes merupakan metode klasifikasi populer dan masuk dalam sepuluh algoritma terbaik dalam data mining, algoritma ini juga dikenal dengan nama Idiot's Bayes, Simple Bayes, dan Independence Bayes[11]

Klasifikasi Naive Bayes adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Klasifikasi bayesian didasarkan pada teorema Bayes, diambil dari nama seorang ahli matematika yang juga menteri Prebysterian Inggris, Thomas Bayes (1702-1761) [12].

Klasifikasi bayesian memiliki kemampuan klasifikasi serupa dengan decision tree dan neural network [13].

$$P(x|y) = \frac{P(y|X) P(x)}{P(y)}$$

(Persamaan 3)

keterangan :

y : data dengan kelas yang belum diketahui

x : hipotesis data y merupakan suatu kelas spesifik

$P(x | y)$  : probabilitas hipotesis x berdasar kondisi y  
(posteriori probability)

$P(x)$  : probabilitas hipotesis x (prior probability)

$P(y | x)$  : probabilitas y berdasarkan kondisi pada hipotesis x

$P(y)$  : probabilitas dari y

#### D. Data Mining

Data Mining merupakan teknologi yang sangat berguna untuk membantu baik itu individu atau perusahaan tertentu dalam menemukan informasi yang sangat penting dari gudang data mereka. Data mining adalah perpaduan dari ilmu statistik, kecerdasan buatan, dan penelitian bidang database [14]. Data mining didefinisikan sebagai proses tentang memecahkan masalah dengan menganalisis data yang berada dalam database [15].

Nama data mining berasal dari kemiripan antara pencarian informasi yang bernilai dari database yang besar dengan menambang sebuah gunung untuk sesuatu yang bernilai [16].

Data mining didefinisikan sebagai proses pengenalan pola dalam data. Data mining, sering disebut *knowledge discovery in database* (KDD), adalah suatu kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menentukan keteraturan, pola atau hubungan dalam set data berukuran besar [15]. Keluaran dari data mining ini biasa dipakai untuk memperbaiki dan membantu pengambilan keputusan di masadepan [17].

### III. METODE PENELITIAN

#### A. Analisis Masalah

Proses analisis dan prediksi dari kartu poker ini menggunakan kombinasi dari kartu yang ada ditangan pemain dengan kartu yang telah diletakkan. Dari hasil kombinasi tersebut maka dapat ditentukan apakah pemain itu berhasil menang atau tidak.

Ada beberapa kombinasi dari pola kartu poker yang dapat terjadi, dengan contoh seperti dibawah ini.



Gambar 1. Kombinasi Kartu Poker

Penentuan peringkat dari pola poker tersebut dimulai dari *royal flush* hingga *high card* dimana posisi atas memiliki nilai paling besar. Penelitian ini menggunakan dataset yang diambil dari situs database <https://archive.ics.uci.edu/ml/datasets/Poker+Hand>. Dataset tersebut memiliki 1.025.010 total data dengan rincian 1.000.000 data berupa testing data dan 25.010 data berupa training data, dan setiap data tersebut memiliki 11 atribut.

Adapun probabilitas kombinasi kartu poker itu antara lain sebagai berikut.

Tabel 1. Probabilitas Kartu Poker

Pola Poker	Kombinasi Tipe (Suit)	Kemungkinan	Jumlah Kombinasi
Royal Flush	4	0.00000154	480
Straight Flush	36	0.00001385	4320
Four of a kind	624	0.0002401	74880
Full House	3744	0.00144058	449280
Flush	5108	0.0019654	612960
Straight	10200	0.00392464	1224000
Three of a kind	54912	0.02112845	6589440
Two Pairs	123552	0.04753902	14826240
One Pair	1098240	0.42256903	131788800
Nothing / High Card	130540	0.50117739	156304800
Total	2598960	1.0	311875200

### IV. HASIL DAN PEMBAHASAN

Adapun pengujian metode yang dilakukan dalam penelitian menggunakan aplikasi WEKA dengan sampel sebanyak 25.010 dan 11 atribut dengan metode pengujian menggunakan, Algoritma C4.5, Decision Stump, Naive Bayes, Bayesian Network, dan Random Forest. Masing-masing metode itu menggunakan *cross-validation* sebanyak sepuluh (10) kali. Adapun rincian dari masing-masing eksperimen itu sebagai berikut.

## A. Eksperimen Algoritma C4.5

```
Time taken to build model: 0.07 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      14390      57.537 %
Incorrectly Classified Instances    10620      42.463 %
Kappa statistic                    0.1954
Mean absolute error                0.0996
Root mean squared error            0.2463
Relative absolute error             86.7951 %
Root relative squared error        103.345 %
Total Number of Instances         25010

==== Detailed Accuracy By Class ====
```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0,808	0,555	0,552	0,808	0,694	0,272	0,656	0,400	0	
0,403	0,242	0,550	0,403	0,465	0,172	0,589	0,498	1	
0,016	0,005	0,129	0,016	0,028	0,029	0,657	0,993	2	
0,004	0,001	0,065	0,004	0,007	0,011	0,645	0,042	3	
0,011	0,001	0,050	0,011	0,018	0,022	0,569	0,006	4	
0,000	0,000	0,000	0,000	0,000	-0,001	0,534	0,002	5	
0,000	0,000	0,000	0,000	0,000	-0,000	0,623	0,006	6	
0,000	0,000	?	0,000	?	?	0,491	0,000	7	
0,000	0,000	?	0,000	?	?	0,496	0,000	8	
0,000	0,000	?	0,000	?	?	0,498	0,000	9	
Weighted Avg.	0,575	0,380	?	0,575	?	?	0,626	0,516	

```
==== Confusion Matrix ====
```

a	b	c	d	e	f	g	h	i	j	<-- classified as
10100	2349	25	7	8	3	0	0	0	0	a = 0
6207	4268	92	20	8	3	1	0	0	0	b = 1
461	723	19	2	0	0	1	0	0	0	c = 2
165	337	9	2	0	0	0	0	0	0	d = 3
60	31	1	0	1	0	0	0	0	0	e = 4
45	9	0	0	0	0	0	0	0	0	f = 5
5	30	1	0	0	0	0	0	0	0	g = 6
0	6	0	0	0	0	0	0	0	0	h = 7
4	0	0	0	1	0	0	0	0	0	i = 8
4	0	0	0	1	0	0	0	0	0	j = 9

Gambar 2. Eksperimen C4.5

Dari gambar pengujian algoritma C4.5 diatas dapat dilihat metode ini memiliki kecepatan proses 0.07 detik dengan tingkat akurasi sebesar 57,537%.

## B. Eksperimen Algoritma Decision Stump

```
Time taken to build model: 0.01 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      12493      49.952 %
Incorrectly Classified Instances    12517      50.048 %
Kappa statistic                    0.2384
Mean absolute error                0.1136
Root mean squared error            0.2384
Relative absolute error             59.9755 %
Root relative squared error        100.0029 %
Total Number of Instances         25010

==== Detailed Accuracy By Class ====
```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
1,000	1,000	0,500	1,000	0,666	?	0,498	0,498	0	
0,000	0,000	?	0,000	?	?	0,496	0,420	1	
0,000	0,000	?	0,000	?	?	0,494	0,048	2	
0,000	0,000	?	0,000	?	?	0,499	0,020	3	
0,000	0,000	?	0,000	?	?	0,517	0,004	4	
0,000	0,000	?	0,000	?	?	0,420	0,002	5	
0,000	0,000	?	0,000	?	?	0,517	0,001	6	
0,000	0,000	?	0,000	?	?	0,215	0,000	7	
0,000	0,000	?	0,000	?	?	0,292	0,000	8	
0,000	0,000	?	0,000	?	?	0,300	0,000	9	
Weighted Avg.	0,500	0,500	?	0,500	?	?	0,497	0,429	

```
==== Confusion Matrix ====
```

a	b	c	d	e	f	g	h	i	j	<-- classified as
12493	0	0	0	0	0	0	0	0	0	a = 0
10599	0	0	0	0	0	0	0	0	0	b = 1
1206	0	0	0	0	0	0	0	0	0	c = 2
513	0	0	0	0	0	0	0	0	0	d = 3
93	0	0	0	0	0	0	0	0	0	e = 4
54	0	0	0	0	0	0	0	0	0	f = 5
36	0	0	0	0	0	0	0	0	0	g = 6
6	0	0	0	0	0	0	0	0	0	h = 7
5	0	0	0	0	0	0	0	0	0	i = 8
5	0	0	0	0	0	0	0	0	0	j = 9

Gambar 3. Eksperimen Decision Stump

Dari gambar pengujian algoritma Decision Stump diatas dapat dilihat metode ini memiliki kecepatan proses 0.01 detik dengan tingkat akurasi sebesar 49,952%.

## C. Eksperimen Algoritma Naive Bayes

```
Time taken to build model: 0 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      12271      49.0644 %
Incorrectly Classified Instances    12739      50.9356 %
Kappa statistic                    -0.0064
Mean absolute error                0.1136
Root mean squared error            0.2387
Relative absolute error             59.9977 %
Root relative squared error        100.1549 %
Total Number of Instances         25010

==== Detailed Accuracy By Class ====
```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0,921	0,929	0,497	0,921	0,646	-0,015	0,503	0,502	0	
0,072	0,077	0,408	0,072	0,122	-0,009	0,497	0,419	1	
0,000	0,000	?	0,000	?	?	0,508	0,051	2	
0,000	0,000	?	0,000	?	?	0,474	0,022	3	
0,000	0,000	?	0,000	?	?	0,667	0,007	4	
0,000	0,000	?	0,000	?	?	0,482	0,004	5	
0,000	0,000	?	0,000	?	?	0,560	0,002	6	
0,000	0,000	?	0,000	?	?	0,715	0,001	7	
0,000	0,000	?	0,000	?	?	0,164	0,000	8	
0,000	0,000	?	0,000	?	?	0,381	0,000	9	
Weighted Avg.	0,491	0,497	?	0,491	?	?	0,501	0,431	

```
==== Confusion Matrix ====
```

a	b	c	d	e	f	g	h	i	j	<-- classified as
11510	983	0	0	0	0	0	0	0	0	a = 0
9838	761	0	0	0	0	0	0	0	0	b = 1
1129	77	0	0	0	0	0	0	0	0	c = 2
480	33	0	0	0	0	0	0	0	0	d = 3
87	6	0	0	0	0	0	0	0	0	e = 4
52	2	0	0	0	0	0	0	0	0	f = 5
33	3	0	0	0	0	0	0	0	0	g = 6
6	0	0	0	0	0	0	0	0	0	h = 7
4	1	0	0	0	0	0	0	0	0	i = 8
4	1	0	0	0	0	0	0	0	0	j = 9

Gambar 4. Eksperimen Naive Bayes

Dari gambar pengujian algoritma Naive Bayes diatas dapat dilihat metode ini memiliki kecepatan proses 0 detik dengan tingkat akurasi sebesar 49,0644%.

## D. Eksperimen Algoritma Bayesian Network

```
Time taken to build model: 0.08 seconds

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      12274      49.0764 %
Incorrectly Classified Instances    12736      50.9236 %
Kappa statistic                    -0.0061
Mean absolute error                0.1136
Root mean squared error            0.2387
Relative absolute error             59.9803 %
Root relative squared error        100.1553 %
Total Number of Instances         25010

==== Detailed Accuracy By Class ====
```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0,921	0,929	0,497	0,921	0,646	-0,015	0,504	0,502	0	
0,072	0,077	0,409	0,072	0,123	-0,009	0,496	0,419	1	
0,000	0,000	?	0,000	?	?	0,508	0,051	2	
0,000	0,000	?	0,000	?	?	0,474	0,022	3	
0,000	0,000	?	0,000	?	?	0,671	0,007	4	
0,000	0,000	?	0,000	?	?	0,487	0,004	5	
0,000	0,000	?	0,000	?	?	0,576	0,002	6	
0,000	0,000	?	0,000	?	?	0,698	0,001	7	
0,000	0,000	?	0,000	?	?	0,172	0,000	8	
0,000	0,000	?	0,000	?	?	0,369	0,000	9	
Weighted Avg.	0,491	0,497	?	0,491	?	?	0,501	0,431	

```
==== Confusion Matrix ====
```

a	b	c	d	e	f	g	h	i	j	<-- classified as
11508	985	0	0	0	0	0	0	0	0	a = 0
9833	766	0	0	0	0	0	0	0	0	b = 1
1129	77	0	0	0	0	0	0	0	0	c = 2
480	33	0	0	0	0	0	0	0	0	d = 3
87	6	0	0	0	0	0	0	0	0	e = 4
52	2	0	0	0	0	0	0	0	0	f = 5
33	3	0	0	0	0	0	0	0	0	g = 6
6	0	0	0	0	0	0	0	0	0	h = 7
4	1	0	0	0	0	0	0	0	0	i = 8
4	1	0	0	0	0	0	0	0	0	j = 9

Gambar 5. Eksperimen Bayesian Network

Dari gambar pengujian algoritma Bayesian Network diatas dapat dilihat metode ini memiliki kecepatan proses 0.08 detik dengan tingkat akurasi sebesar 48,0764%.



## E. Eksperimen Algoritma Random Forest

Time taken to build model: 2.26 seconds

==== Stratified cross-validation ====

==== Summary ====

Correctly Classified Instances1857974.2863 %

Incorrectly Classified Instances643125.7137 %

Kappa statistic0.5152

Mean absolute error0.0949

Root mean squared error0.2065

Relative absolute error83.5421 %

Root relative squared error86.657 %

Total Number of Instances25010

==== Detailed Accuracy By Class ====

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
0,925	0,300	0,755	0,925	0,831	0,641	0,948	0,946	0	
0,659	0,185	0,724	0,659	0,690	0,481	0,780	0,635	1	
0,017	0,000	0,870	0,017	0,033	0,116	0,850	0,274	2	
0,023	0,000	0,923	0,023	0,046	0,145	0,951	0,446	3	
0,000	0,000	?	0,000	?	?	0,776	0,030	4	
0,000	0,000	0,000	0,000	0,000	-0,000	0,580	0,003	5	
0,000	0,000	?	0,000	?	?	0,824	0,027	6	
0,000	0,000	?	0,000	?	?	0,654	0,004	7	
0,000	0,000	0,000	0,000	0,000	-0,000	0,484	0,000	8	
0,000	0,000	?	0,000	?	?	0,594	0,001	9	
Weighted Avg.	0,743	0,228	?	0,743	?	?	0,870	0,764	

==== Confusion Matrix ====

a	b	c	d	e	f	g	h	i	j	<-- classified as
11558	934	0	0	0	1	0	0	0	0	a = 0
3608	6989	2	0	0	0	0	0	0	0	b = 1
7	1179	20	0	0	0	0	0	0	0	c = 2
0	500	1	12	0	0	0	0	0	0	d = 3
93	9	0	0	0	0	0	1	0	0	e = 4
51	3	0	0	0	0	0	0	0	0	f = 5
0	36	0	0	0	0	0	0	0	0	g = 6
0	5	0	1	0	0	0	0	0	0	h = 7
5	0	0	0	0	0	0	0	0	0	i = 8
3	2	0	0	0	0	0	0	0	0	j = 9

Gambar 6. Eksperimen Random Forest

Dari gambar pengujian algoritma Random Forest diatas dapat dilihat metode ini memiliki kecepatan proses 2,26 detik dengan tingkat akurasi sebesar 74,2863%.

Dari lima eksperimen yang telah dilakukan diatas didapatkan peringkat dari metode itu sebagai berikut.

```

Tester: weka.experiment.PairedCorrectedTester -O 4,5,6 -D 1 -R 2 -S 0.05 -V -result-matrix "weka.experiment.ResultMatrixPlainTest -mean-per
Analyzing: Percent_correct
Dataset: 1
Results: 5
Confidence: 0.05 (two tailed)
Sorted by:
Date: 25/10/19 15:56

```

```

>C > < Resultset
4 4 0 trees.RandomForest "-P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1" 114839470751428498
2 3 1 trees.J48 "-C 0.25 -M 2" -21773216939364444
0 2 2 trees.DecisionStump "" 161838453550391
-3 0 3 bayes.NaiveBayes "" 5995232201795497455
-3 0 3 bayes.Bayesian "-O -Q bayes.net.search.local.k2 -- -P 1 -S BAYES -E bayes.net.estimate.SimpleEstimator -- -A 0.5" 74603743250775954

```

Gambar 7. Peringkat Eksperimen

## V. KESIMPULAN

Didalam penelitian ini dilakukan eksperimen pada lima algoritma data mining untuk prediksi pola kartu poker. Dari eksperimen itu metode Random Forest menempati peringkat pertama yang disusul dengan metode C4.5, Decision Stump, Naive Bayes dan Bayesian Network. Tetapi meskipun metode Random Forest tersebut memiliki peringkat pertama dengan akurasi tertinggi tetapi metode tersebut membutuhkan waktu yang lama sebesar 2,26 detik, jauh jika dibandingkan dengan metode lain yang hanya menggunakan waktu 0 sampai 0.08 detik.

## Daftar Pustaka

- [1] T. L. MacKay, N. Bard, M. Bowling, and D. C. Hodgins, "Do pokers players know how good they are? Accuracy of poker skill estimation in online and offline players," *Comput. Human Behav.*, vol. 31, no. 1, pp. 419–424, 2014.
- [2] B. Zaman, K. Geurden, R. De Cock, B. De Schutter, and V. Vanden Abele, "Motivation profiles of online Poker players and the role of interface preferences: A ladder study among amateur and (semi-) professionals," *Comput. Human Behav.*, vol. 39, pp. 154–164, 2014.
- [3] M. Ponsen, K. Tuyls, M. Kaisers, and J. Ramon, "An evolutionary game-theoretic analysis of poker strategies," *Entertain. Comput.*,

vol. 1, no. 1, pp. 39–45, 2009.

- [4] G. Ambekar, T. Chikane, S. Sheth, A. Sable, and K. Ghag, "Anticipation of winning probability in poker using data mining," in *IEEE International Conference on Computer Communication and Control, IC4 2015*, 2016.
- [5] F. H. Chen, D. J. Chi, and Y. C. Wang, "Detecting biotechnology industry's earnings management using Bayesian network, principal component analysis, back propagation neural network, and decision tree," *Econ. Model.*, vol. 46, pp. 1–10, 2015.
- [6] M. Vazifteh, M. H. Moattar, and M. Jalali, "A hybrid Bayesian network and tensor factorization approach for missing value imputation to improve breast cancer recurrence prediction," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 31, no. 2, pp. 175–184, 2019.
- [7] H. Elouedi, W. Meliani, Z. Elouedi, and N. Ben Amor, "A hybrid approach based on decision trees and clustering for breast cancer classification," *6th Int. Conf. Soft Comput. Pattern Recognition, SoCPaR 2014*, pp. 226–231, 2015.
- [8] C. Xiang, P. C. Yong, and L. S. Meng, "Design of multiple-level hybrid classifier for intrusion detection system using Bayesian clustering and decision trees," *Pattern Recognit. Lett.*, vol. 29, no. 7, pp. 918–924, 2008.
- [9] P. Ruangthong and S. Jaiyen, "Hybrid ensembles of decision trees and Bayesian network for class imbalance problem," *2016 8th Int. Conf. Knowl. Smart Technol. KST 2016*, pp. 39–42, 2016.
- [10] S. Jabin, "Poker hand classification," pp. 269–273, 2016.
- [11] M. Bramer, *Principles of Data Mining*, vol. 98, no. 461. London: Springer London, 2016.
- [12] D. T. Larose and C. D. Larose, *Discovering Knowledge in Data*, vol. 190, no. 5. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2014.
- [13] O. Maimon and L. Rokach, "Introduction to Knowledge Discovery and Data Mining," in *Data Mining and Knowledge Discovery Handbook*, 2009.
- [14] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. 2012.
- [15] I. H. Witten, E. Frank, and M. a. Hall, *Data Mining: Practical Machine Learning Tools and Techniques, Third Edition*. 2011.
- [16] S. Sumathi and S. N. Sivanandam, *Introduction to Data Mining and its Applications*. 2006.
- [17] S. Budi, *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. 2007.